A hierarchical image segmentation algorithm based on an observation scale

Silvio Jamil Ferzoli Guimarães, Jean Cousty, Yukiko Kenmochi, Laurent Najman

A Hierarchical Image Segmentation Algorithm Based on an Observation Scale*

Silvio Jamil F. Guimarães[1,2], Jean Cousty[2], Yukiko Kenmochi[2], and Laurent Najman[2]

[1] PUC Minas - ICEI - DCC - VIPLAB
sjamil@pucminas.br
[2] Université Paris-Est, LIGM, ESIEE - UPEMLV - CNRS
{j.cousty,y.kenmochi,l.najman}@esiee.fr

Abstract. Hierarchical image segmentation provides a region-oriented scale-space, *i.e.*, a set of image segmentations at different detail levels in which the segmentations at finer levels are nested with respect to those at coarser levels. Most image segmentation algorithms, such as region merging algorithms, rely on a criterion for merging that does not lead to a hierarchy. In addition, for image segmentation, the tuning of the parameters can be difficult. In this work, we propose a hierarchical graph based image segmentation relying on a criterion popularized by Felzenszwalb and Huttenlocher. Quantitative and qualitative assessments of the method on Berkeley image database shows efficiency, ease of use and robustness of our method.

Keywords: hierarchical segmentation, edge-weighted graph, saliency map.

1 Introduction

Image segmentation is the process of grouping perceptually similar pixels into regions. A hierarchical image segmentation is a set of image segmentations at different detail levels in which the segmentations at coarser detail levels can be produced from simple merges of regions from segmentations at finer detail levels. Therefore, the segmentations at finer levels are nested with respect to those at coarser levels. Hierarchical methods have the interesting property of preserving spatial and neighboring information among segmented regions. Here, we propose a hierarchical image segmentation in the framework of edge-weighted graphs, where the image is equipped with an adjacency graph and the cost of an edge is given by a dissimilarity between two points of the image.

Any hierarchy can be represented with a minimum spanning tree. The first appearance of this tree in pattern recognition dates back to the seminal work

* The authors are grateful to FAPEMIG and CAPES, which are Brazilian research funding agencies, and also to Agence Nationale de la Recherche through contract ANR-2010-BLAN-0205-03 KIDICO, which is a French research funding agency.

G.L. Gimel' farb et al. (Eds.): SSPR & SPR 2012, LNCS 7626, pp. 116–125, 2012.

of Zahn [1]. Lately, its use for image segmentation was introduced by Morris *et al.* [2] in 1986 and popularized in 2004 by Felzenszwalb and Huttenlocher [3]. However the region-merging method [3] does not provide a hierarchy. In [4,5], it was studied some optimality properties of hierarchical segmentations. Considering that, for a given image, one can tune the parameters of the well-known method [3] for obtaining a reasonable segmentation of this image. We provide in this paper a hierarchical version of this method that removes the need for parameter tuning.

The algorithm of [3] is the following. First, a minimum spanning tree (MST) is computed, and all the decisions are taken on this tree. For each edge linking two vertices x and y, following a non-decreasing order of their weights, the following steps are performed:

(i) Find the region X that contains x.
(ii) Find the region Y that contains y.
(iii) Merge X and Y according to a certain criterion.

The criterion for region-merging in [3] measures the evidence for a boundary between two regions by comparing two quantities: one based on intensity differences across the boundary, and the other based on intensity differences between neighboring pixels within each region. More precisely, in step (iii), in order to know whether two regions must be merged, two measures are considered. The *internal difference* $Int(X)$ of a region X is the highest edge weight among all the edges linking two vertices of X in the MST. The *difference* $Diff(X, Y)$ between two neighboring regions X and Y is the smallest edge weight among all the edges that link X to Y. Then, two regions X and Y are merged when:

$$Diff(X, Y) \leq \min\{Int(X) + \frac{k}{|X|}, Int(Y) + \frac{k}{|Y|}\} \tag{1}$$

where k is a parameter allowing to prevent the merging of large regions (*i.e.*, larger k forces smaller regions to be merged).

The merging criterion defined by Eq. (1) depends on the scale k at which the regions X and Y are observed. More precisely, let us consider the *(observation) scale* $S_Y(X)$ *of* X *relative to* Y as a measure based on the difference between X and Y, on the internal difference of X and on the size $|X|$ of X:

$$S_Y(X) = (Diff(X, Y) - Int(X)) \times |X|. \tag{2}$$

Then, the *scale* $S(X, Y)$ is simply defined as:

$$S(X, Y) = \max(S_Y(X), S_X(Y)). \tag{3}$$

Thanks to this notion of a scale, Eq. (1) can be written as:

$$k \geq S(X, Y). \tag{4}$$

In other words, Eq.(4) states that the neighboring regions X and Y merge when their scale is less than the threshold parameter k.

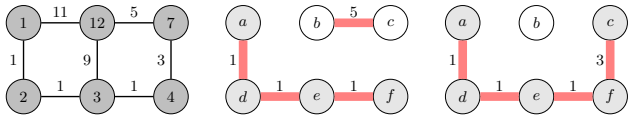
Even if the image segmentation results obtained by the method proposed in [3] are interesting, the user faces two major issues:

(a) Original (b) $k = 7500$ (8) (c) $k = 9000$ (14)

Fig. 1. A real example illustrating the violation of the causality principle by [3]: the number of regions (in parentheses) increases from 8 to 14, instead of decreasing when the so-called "scale of observation" increases

- first, the number of regions may increase when the parameter k increases. This should not be possible if k was a true scale of observation: indeed, it violates the *causality principle* of multi-scale analysis, that states in our case [6] that a contour present at a scale k_1 should be present at any scale $k_2 < k_1$. Such unexpected behaviour of missing causality principle is demonstrated on Fig. 1.
- Second, even when the number of regions decreases, contours are not stable: they can move when the parameter k varies, violating a *location principle*. Such a situation is illustrated on Fig. 2.

Given these two issues, the tuning of the parameters of [3] is a difficult task.



(a) Original image (b) $k = 5$ (c) $k = 8$

Fig. 2. An example illustrating the violation of the location principle by [3]: the contours are unstable from one "scale" to another

Following [6], we believe that, in order for k to be a true scale-parameter, we have to satisfy both the causality principle and the location principle, which leads to work with a hierarchy of segmentations. Reference [7] is the first to propose an algorithm producing a hierarchy of segmentations based on [3]. However, this method is an iterative version of [3] that uses a threshold function, and requires a tuning of the threshold parameter.

The main result of this paper is an efficient hierarchical image segmentation algorithm based on the dissimilarity measure of [3]. Our algorithm has a computational cost similar to [3], but provides all scales of observations instead of only one segmentation level. As it is a hierarchy, the result of our algorithm satisfies both the locality principle and the causality principle. Namely, and in contrast with [3], the number of regions is decreasing when the scale parameter increases, and the contours do not move from one scale to another.

Figure 3 illustrates the results obtained by applying our method to the same image of Fig. 1(a), with segmentations at two different scales of observations, as

(a) Saliency map (b) $k = 1000$ (22) (c) $k = 5000$ (6)

Fig. 3. A real example illustrating the saliency map of Fig. 1(a) computed with our approach. We display in (b) and (c) two image segmentations extracted from the hierarchy at scales 1000 and 5000, together with their numbers of regions (in parentheses).

well as a saliency map [8,4,5] (a map indicating the disparition level of contours and whose thresholds give the set of all segmentations).

This work is organized as follows. In Section 2, we present our hierarchical method for color image segmentation. Some experimental results performed on Berkeley image database are given in Section 3. Finally, in Section 4, some conclusions are drawn and further works are discussed.

2 A Hierarchical Graph Based Image Segmentation

In this section, we describe our method to compute a hierarchy of partitions based on observation scales as defined by Eq. (3). Let us first recall some important notions for handling hierarchies [2,4,5].

To every tree T spanning the set V of the image pixels, to every map $w : E \to \mathbb{N}$ that weights the edges of T and to every threshold $\lambda \in \mathbb{N}$, one may associate the partition $\mathcal{P}^w_\lambda$ of V induced by the connected components of the graph made from V and the edges of weight below λ. It is well known [2,5] that for any two values λ_1 and λ_2 such that $\lambda_1 \geq \lambda_2$, the partitions $\mathcal{P}^w_{\lambda_1}$ and $\mathcal{P}^w_{\lambda_2}$ are *nested* and $\mathcal{P}^w_{\lambda_1}$ is *coarser* than $\mathcal{P}^w_{\lambda_2}$. Hence, the set $\mathcal{H}^w = \{\mathcal{P}^w_\lambda \mid \lambda \in \mathbb{N}\}$ is a *hierarchy of partitions induced by the weight map* w.

Our algorithm does not explicitly produce a hierarchy of partitions, but instead produces a weight map L (scales of observations) from which the desired hierarchy $\mathcal{H}^L$ can be inferred on a given T. It starts from a minimum spanning tree T of the edge-weighted graph built from the image. In order to compute the scale $L(e)$ associated with each edge of T, our method iteratively considers the edges of T in a non-decreasing order of their original weights w. For every edge e, the new weight map $L(e)$ is initialized to ∞; then for each edge e linking two vertices x and y the following steps are performed:

(i) Find the the region X of $\mathcal{P}^w_{w(e)}$ that contains x.
(ii) Find the the region Y of $\mathcal{P}^w_{w(e)}$ that contains y.
(iii) Compute the hierarchical observation scale $L(e)$.

At step (iii), the *hierarchical scale* $S'_Y(X)$ of X relative to Y is needed to obtain the value $L(e)$. Intuitively, $S'_Y(X)$ is the lowest observation scale at which some sub-region of X, namely X^*, will be merged to Y. More precisely, using an internal parameter v, this scale is computed as follows:

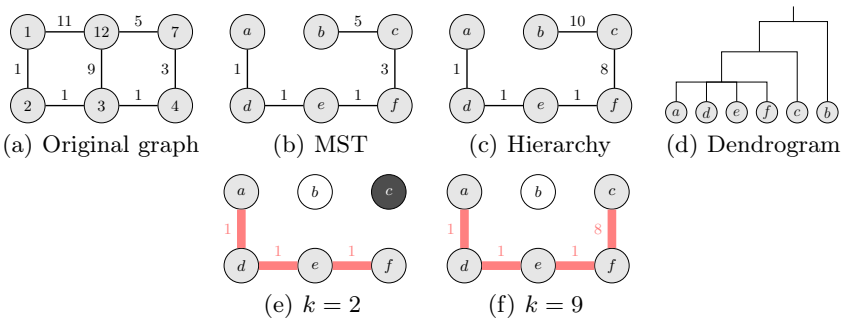


Fig. 4. Example of hierarchical image segmentations. In contrast to example in Fig. 2, the contours are stable from a scale to another, providing a hierarchy.

(1) Initialize the value of v to ∞.
(2) Decrement the value of v by 1.
(3) Find the the region X^* of $\mathcal{P}_v^L$ that contains x.
(4) Repeat steps 2 and 3 while $S_Y(X^*) < v$
(5) Set $S'_Y(X) = v$.

With the appropriate changes, the same algorithm allows $S'_X(Y)$ to be computed. Then, the hierarchical scale $L(e)$ is simply set to:

$$L(e) = \max\{S'_Y(X), S'_X(Y)\}. \tag{5}$$

Figure 4 illustrates the result of our method on a pedagogical example. Starting from the graph of Fig. 4(a), our method produces the hierarchical observation scales depicted in Fig. 4(c). As for the method of [3], our algorithm only considers the edges of the minimum spanning tree (see Fig. 4(b)). The whole hierarchy is depicted as a dendrogram in Fig. 4(d), whereas two levels of the hierarchy (at scales 2 and 9) are shown in Fig. 4(e) and (f).

2.1 Implementation Issues

To efficiently implement our method, we use some data structures similar to the ones proposed in [5]; in particular, the management of the collection of partitions is due to Tarjan's union find and Fredman and Tarjan's Fibonnacci heaps. Furthermore, we made some algorithmic optimizations to speed up the computations of the observation scales. In order to illustrate an example of computation time, we implemented all our algorithm in C++ on a standard single CPU computer under windows Vista, we run it in a Intel Core 2 Duo, 4GB. For the image illustrated in Fig. 1(a) (with size 321x481), the hierarchy is computed in 2.7 seconds, and the method proposed in [3] spent 1.3 seconds.

3 Experimental Results

In this section, we present a quantitative and a qualitative assessments in order to better compare our method to the method proposed in [3] (called method FH

Fig. 5. An example of a hierarchical image segmentations of a synthetic image containing three perceptually big regions. The saliency map of the image (a) is shown in (b). The number of regions of the segmented images is written under each figure.

Fig. 6. Top row: some images of the Berkeley database [9]. Middle row: saliency maps of these images according to our hierarchical method. The numbers of scales of these hierarchies are (a) 240, (b) 443, (c) 405 and (d) 429. Bottom row: according to our subjective judgment, the best segmentations extracted from the hierarchies. The numbers of regions are (a) 3, (b) 18, (c) 6 and (d) 16.

Table 1. Performances of our method and the method FH [3] using two different measures: Ground-truth Covering (GT Covering) and Probabilistic Rand Index. The presented scores are optimal considering a constant scale parameter for the whole dataset (ODS) and a scale parameter varying for each image (OIS). See [9] for more details on the evaluation method.

Area	GT Covering				Prob. Rand. Index			
	ODS		OIS		ODS		OIS	
	Ours	FH	Ours	FH	Ours	FH	Ours	FH
20	0.42	0.43	0.52	0.52	0.75	0.75	0.81	0.79
50	0.44	0.43	0.52	0.52	0.76	0.75	0.81	0.79
500	0.46	0.43	0.53	0.53	0.76	0.76	0.81	0.79
1000	0.46	0.44	0.53	0.53	0.76	0.76	0.80	0.80
1500	0.46	0.44	0.52	0.54	0.76	0.76	0.80	0.80

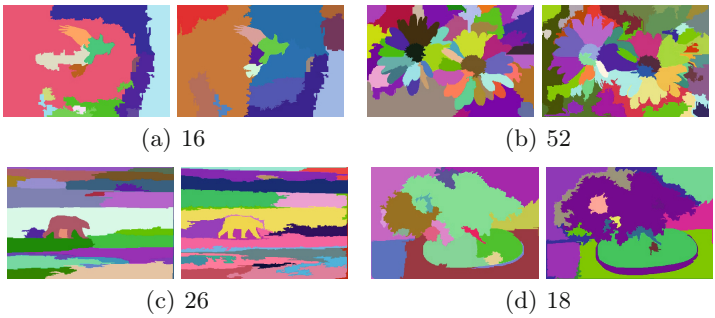


Fig. 7. Comparison between our method and the method FH [3]. For each pair of images, the right image shows the best result (according to our judgment and our experiments) from [3] and the left image shows a segmentation extracted from our hierarchical result, with the same number of regions.

hereafter). The former is based on evaluation framework proposed in [9], and the later one is based on three experiments in which we tune the parameters to (visually) evaluate the quality of the segmentations. A major difficulty of experiments is the design of an adequate edge-cost, well adapted to the content to be segmented. A practical solution is to use some dissimilarity functions, and many different functions are used in the literature. In this work, the underlying graph is the one induced by the 4-adjacency pixel relation, where the edges are weighted by a simple color gradient computed by the Euclidean distance in the RGB space. Before presenting the quantitative and qualitative assessments, we illustrate some results of our method.

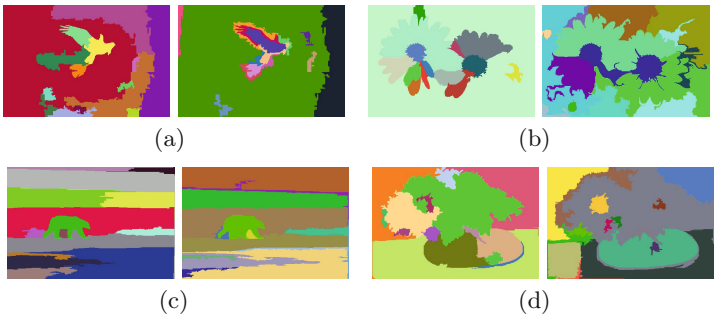


Fig. 8. Examples of image segmentation where the number of regions has been set to 15. For each pair of images, the left one shows a segmentation extracted from our hierarchy; and the right one shows the result obtained with [3] by varying the parameter k until the desired number of regions is found.

(a) (b) (c)

Fig. 9. Examples of segmentations for images corrupted by a random salt noise. The corrupted images (at different levels - 70% and 90%) are shown on the first column. The results of our method and [3] are illustrated in the second and third columns, respectively.

In Fig. 5, we present some results on an artificial image containing three perceptually large regions. With this example, one can easily verify the hierarchical property of our method by looking at the segmentations at scales *resp.* 1000, 2000, 5000, 140000 and 224000 (*resp.* Fig. 5(c), (d), (e), (f) and (g)). Since the resulting segmentations are nested, the whole hierarchy can be presented in a saliency map (see Fig. 5(b)). Figure 6 illustrates the performance of our method when applied to some images of the Berkeley's database [9]. Note that, as in [3], an area filtering is applied to eliminate small regions (smaller than 500 pixels).

In the sequel, we present the quantitative assessment followed by the qualitative one. Table 1 assesses the equivalent performances of our method and of the method FH [3], according to the evaluation framework proposed in [9], in terms of Ground-truth Covering and Probabilistic Rand Index, when applied on 200 test images of the Berkeley's database [9]. For this experiment, an area filtering is applied to eliminate small regions varying from 20 to 1500 pixels.

For the qualitative assessment, we made three experiments. First, we try to set the reasonable parameter for [3], *i.e.* the parameter that produces the best

(subjective) visual result (Fig. 7). We can compare this result with the segmentation result at a scale in our hierarchy such that it contains the same number of regions that of [3] (Fig. 7). In a second experiment, we fixed the number of regions to 15 for all images, and tune the parameter for [3] to obtain this number of regions. For our method, we use breadth-first traversal in the hierarchy (tree structure) to find the scale that givens 15 regions. We compare those segmentations in Fig. 8. The last illustration (Fig. 9) is designed to assess the robustness to random impulse noise. From these experiments, we observe that in general, our method produces "objects" (or regions) better defined with respect to the results obtained by the method FH. Moreover, the contours are stable, *i.e.*, the contours do not move from one scale to another, and the number of regions is decreasing when the scale parameter increases.

4 Conclusions

This paper proposes an efficient hierarchical segmentation method based on the observation scales of [3]. In contrast to [3], our method produces the complete set of segmentations at every scale, and satisfies both the causality and location principle defined by [6]. An important practical consequence of these properties is to ease the selection of a scale level adapted to a particular task. We assess our method and the method of [3] on the Berkeley database following the methodology introduced in [9]. We visually assessed our method on some real images by comparing our segmentations to those of [3]. From theses quantitative and qualitative assessments, the produced segmentations are promising, in particular w.r.t. robustness. As future work, we will investigate using more information into the definition of observation scale as well as learning which information is pertinent for a given practical task. Moreover, we will investigate theoretical properties of our method.

References

1. Zahn, C.T.: Graph-theoretical methods for detecting and describing gestalt clusters. IEEE Trans. Comput. 20, 68–86 (1971)
2. Morris, O., Lee, M.J., Constantinides, A.: Graph theory for image analysis: an approach based on the shortest spanning tree. Communications, Radar and Signal Processing, IEE Proceedings F 133(2), 146–152 (1986)
3. Felzenszwalb, P.F., Huttenlocher, D.P.: Efficient graph-based image segmentation. IJCV 59, 167–181 (2004)
4. Najman, L.: On the equivalence between hierarchical segmentations and ultrametric watersheds. JMIV 40, 231–247 (2011)
5. Cousty, J., Najman, L.: Incremental Algorithm for Hierarchical Minimum Spanning Forests and Saliency of Watershed Cuts. In: Soille, P., Pesaresi, M., Ouzounis, G.K. (eds.) ISMM 2011. LNCS, vol. 6671, pp. 272–283. Springer, Heidelberg (2011)

6. Guigues, L., Cocquerez, J.P., Men, H.L.: Scale-sets image analysis. IJCV 68(3), 289–317 (2006)
7. Haxhimusa, Y., Kropatsch, W.: Segmentation Graph Hierarchies. In: Fred, A., Caelli, T.M., Duin, R.P.W., Campilho, A.C., de Ridder, D. (eds.) SSPR&SPR 2004. LNCS, vol. 3138, pp. 343–351. Springer, Heidelberg (2004)
8. Najman, L., Schmitt, M.: Geodesic saliency of watershed contours and hierarchical segmentation. PAMI 18(12), 1163–1173 (1996)
9. Arbelaez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. PAMI 33, 898–916 (2011)